
Waiting Strategies for Regular and Emergency Patient Transportation

Guenter Kiechle¹, Karl F. Doerner², Michel Gendreau³, and Richard F. Hartl²

¹ Vienna Technical University, Karlsplatz 13, 1040 Vienna, Austria
guenter.kiechle@salzburgresearch.at

² University of Vienna, Department of Business Administration, Bruenner Strasse 72, 1210 Vienna, Austria {karl.doerner, richard.hartl}@univie.ac.at

³ Centre interuniversitaire de recherche sur les réseaux d'entreprise, la logistique et le transport (CIRRELT), C.P. 6128, succursale Centre-ville, Montréal, Canada H3C 3J7 michelg@crt.umontreal.ca

1 Introduction

Many emergency service providers, especially ambulance departments and companies who provide non-public maintenance services, face the problem that their fleet of vehicles has to provide two different types of services:

1. Cover a certain region and provide immediate service when an emergency occurs;
2. Provide some regular service (e.g., the pick-up and delivery of patients, predetermined service tasks, periodic pick-ups...).

This is the current situation for the largest Austrian emergency service providers (e.g., the Austrian Red Cross), where the same fleet is used to provide both types of services. Dynamic aspects thus directly influence the schedule for the regular service. When an emergency occurs and an ambulance is required, the vehicle with the shortest distance to the emergency is assigned to serve the emergency patient. Therefore, it may happen that an ambulance vehicle that has to carry out a scheduled transport order of a patient, which has not started yet, is used to serve the emergency request and the schedule for the regular services has to be re-optimized and another vehicle has to be reassigned to the regular patient. Ambulances that carry out emergency transport

become available at the hospital after the emergency service and can be then used to carry out regular transport orders. Again, the schedule for regular services has to be re-optimized.

From the perspective of managing the regular services, the objective is minimizing the total traveling distance subject to certain restrictions (e.g., be on time). From the perspective of minimizing the response time for servicing an emergency request, it is necessary to locate and schedule the vehicles in such a way that each possible location where an emergency case may occur can be reached within a given time (see [4]).

These two objectives are not totally contradictory, but they certainly conflict: on the one hand, for the emergency transport requests, one has to account for the fact that vehicles are changing positions and are blocked for some time due to some regular service assignment; on the other hand, when regular orders are assigned to vehicles, it is important to keep a certain coverage level to ensure a satisfactory service of the emergency cases, which may require relocating some of the vehicles.

Emergency service providers thus want to find a robust solution for a specific day of the week in order to minimize routing costs, as well as the response time for servicing emergency patients, taking into account that the two types of the transportation services have to be realized with a single fleet.

Some related work has been published where pickup and delivery requests occur dynamically (see [1, 3, 6]). Note that in our case only emergency orders occur dynamically. The nearest empty vehicle is used to serve the emergency immediately. Since subsequent scheduled regular transport requests on the redirected vehicle are not covered any more, a re-optimization step is initiated.

In order to study different dispatch strategies for the real world patient transportation problem, we need a simple, fast and effective solution procedure. Hence, we implemented a constructive heuristic approach. In the construction phase, we exploit the temporal structure of the requests and use a nearest neighbor measure inspired by [5]. The main challenge is to deal with the dynamic nature of the problem, which implies that new solutions can be calculated in very short time. Every time a new emergency request occurs, the distance information for the empty vehicles has to be updated in order to identify very quickly the next empty vehicle to the emergency request. Then, one has to resolve the remaining problem with one less vehicle available to serve the regular orders. When an emergency order is fulfilled and the patient has been unloaded at the hospital, an additional vehicle is made available

and it can be used to fulfill regular patient transportation orders. When this situation occurs, the schedule for the regular transport orders has also to be re-optimized to take advantage of this additional vehicle to improve the objective function.

To evaluate the solution quality various scenarios have to be calculated. In the evaluation of the different scenarios, vehicle movements and the spatial distribution of empty vehicles are monitored over time. The distance matrix is also updated whenever an emergency request occurs.

2 Problem Description

In our study we consider a combination of two problems

- the Dial-a-Ride Problem (DARP) for regular patient transportation and
- dispatching ambulance vehicles for emergency cases.

In the classical DARP, a set of requests announced beforehand are served from a single depot. These requests have hard time windows and a preferred pickup or delivery time. In the problem of dispatching ambulance vehicles for emergency cases, one must ensure short response times in a dynamic environment. The regular patient transportation problem can be considered to be a variation of DARP with additional real world constraints regarding customer preferences or requirements. A comprehensive description of the DARP is given in [2].

Each transport order or request incorporates a pickup location and a delivery location. For each transportation request a time window and a service or loading time for each pickup and delivery location is given. Concerning time windows, we have two different situations - on the one hand, patients should be picked-up as late as possible from their home when they are being transported to hospitals (outbound request); on the other hand, patients should be picked-up as early as possible when they are transported from the hospital back home (inbound request). Time window violations are not allowed and deviations from the desired pick-up and drop-off times within the specified time window are considered in the objective function.

Time windows for each request are defined either by a desired delivery time for outbound requests or a desired pickup time for inbound requests. The resulting time window is calculated by allowing an early pickup or an early delivery by 30 minutes. The pickup or delivery has to be performed within this time window. We allow also a maximum

ride time for each passenger. This maximum ride time is defined for each customer by calculating the shortest path from the pickup to the delivery location and allowing an additional ride time of 30 minutes. In our computational tests the vehicle capacity is two, therefore the maximum number of patients that can be transported at the same time is two. Waiting is not allowed with a passenger on board.

In our real world problem for the regular patients it is desired to minimize transportation costs and to maximize quality of service for patients. And for the emergency patients it is desired to minimize response time for emergency requests.

In the objective function for the constructive procedure the regular patients (transportation costs, quality of service) are considered whereas the response time minimization is considered in the different waiting strategies. Quality of service for the regular transport orders is measured by the deviation of the preferred delivery time for the outbound request and the deviation of the preferred pickup time for the inbound request respectively.

3 Dynamic aspects

After computing a solution for the problem at hand using one of the proposed solution procedures, improving the coverage criterion is possible without changing the solution structure. More precisely, movements of vehicles may be delayed or pushed forward in time while the sequence of transport requests on each tour remains unchanged.

The distribution is measured at discrete points in time for all empty vehicles, e.g., at equal intervals from the first pick-up to the last drop-off during the day or optimization period. Besides the empty vehicle positions, we use a set of equidistant raster points in the respective area. The covered raster points by the empty vehicles within a certain radius are calculated and used as a proxy for the response time.

Therefore, we distinguish four different waiting strategies to influence movements of empty vehicles on their way from a delivery location to the next pickup location. The first two strategies are static ones called 'early arrival' and 'late arrival'. Early arrival means, that a vehicle departs to its next pickup location immediately after delivering the last patient and waits at the pickup location until the next pickup starts. On the contrary, late arrival means, that a vehicle waits at the last delivery location just as long as possible to arrive at the next pickup location in time.

The last two strategies are dynamic variants of the former ones, which use late arrival or early arrival as default, but may change the strategy for each single request if an individual coverage calculation indicates an improvement. For each single request both coverage values (late and early arrival) are calculated. The strategy with the better global coverage is performed. Note that potential movements of other vehicles while the respective vehicle drives from delivery to pickup location are not considered in this proxy evaluation.

4 Evaluation of the approach

In order to test the solution procedures with a comprehensive set of parameters and emergency instances, a simulation environment for the problem at hand was implemented. The efficiency of the solution procedures was evaluated with real world data from the city of Graz provided by the Austrian Red Cross. We used 15 problem instances with a number of regular transport requests ranging from 152 to 286 and grouped them into three classes of five instances each. The small instances encounter from 152 to 192, the medium instances from 200 to 220 and the large instances from 260 to 286 transport requests. All requests occur after 6:00 a.m. and before 5:00 p.m. and for each outbound request a corresponding inbound request occurs a few hours later.

In the computational studies, we evaluated the impact of different waiting strategies as described above. Hence, we evaluated response times to artificial emergency transport requests for a large number of runs. For each instance and each waiting strategy we measured response times for three randomly distributed emergency requests over 2000 independent runs. In Table 1 we report average response time, maximum response time and coverage for each instance class as well as for each waiting strategy. The improvements of the dynamic strategies compared to the static counterpart are reported in the last two rows of the tables (imp.).

The results show that the dynamic strategies outperform the static ones, whereas early arrival generally performs better than late arrival. The average response time using early arrival strategy can be reduced from 5.04 to 5.01 minutes. Also the maximum response time in the case of early arrivals can be reduced from 7.87 minutes to 7.83 minutes. The improved coverage enables these improvements. By using a dynamic waiting strategy the coverage can be improved e.g. for the early arrival strategy by 1.65 %. Improvements from the static to the dynamic strategies are remarkably higher for late arrival.

instance class	average response time				maximum response time			
	small	medium	large	average	small	medium	large	average
static late arrival	5.38	5.19	5.12	5.23	8.32	8.09	7.93	8.11
static early arrival	5.20	5.00	4.94	5.04	8.07	7.85	7.69	7.87
dynamic late arrival	5.17	5.00	4.97	5.05	8.03	7.83	7.72	7.86
dynamic early arrival	5.14	4.97	4.93	5.01	8.01	7.81	7.67	7.83
imp. late arrival	4.06%	3.65%	2.97%	3.57%	3.49%	3.39%	2.74%	3.21%
imp. early arrival	1.18%	0.54%	0.21%	0.65%	0.78%	0.50%	0.20%	0.50%

instance class	coverage			
	small	medium	large	average
static late arrival	30.0%	32.4%	31.0%	31.2%
static early arrival	30.6%	33.2%	32.2%	32.0%
dynamic late arrival	32.0%	34.4%	33.1%	33.2%
dynamic early arrival	32.2%	34.6%	33.3%	33.4%
imp. late arrival	6.16%	5.81%	6.43%	6.13%
imp. early arrival	4.90%	3.93%	3.56%	4.12%

Table 1. Response times and coverage subject to waiting strategies

Acknowledgements Financial support from the Austrian Science Fund (FWF) under grant #L286-N04 is gratefully acknowledged.

References

1. Attanasio, A., Cordeau, J.-F., Ghiani, G., and Laporte, G. (2004) Parallel Tabu search heuristics for the dynamic multi-vehicle dial-a-ride problem. *Parallel Computing* 30:377–387.
2. Cordeau, J.-F., and Laporte, G. (2003) “The Dial-a-Ride Problem (DARP): Variants modelling issues and algorithms”. *4OR* 1, pp. 89–101.
3. Gendreau, M., Guertin, F., Potvin, J.-Y., and Séguin, R. (2002): “Neighborhood Search heuristic for a Dynamic Vehicle Dispatching Problem with Pickups and Deliveries”. Technical Report, Centre de recherche sur les transports, Université de Montréal. Forthcoming in *Transportation Research C*.
4. Gendreau, M., Laporte, G., and Semet, F. (2001): “A dynamic model and parallel tabu search heuristic for real-time ambulance relocation”. *Parallel Computing* 27, 1641–1653.
5. Jaw J.-J., Odoni, A., Psaraftis, H., Wilson, N. (1986): “A heuristic algorithm for the multi-vehicle advance request dial-a-ride problem with time windows”. *Transportation Research B* 20 (3), pp. 243-257.
6. Mitrović-Minić, S., and Laporte, G. (2004): “Waiting Strategies for the Dynamic Pickup and Delivery Problem with Time Windows”. *Transportation Research B* 38, pp. 635-655.